"Proofs"

(1) Prediction Eqⁿ:-

$$y = a + bx$$

$$\sum y_i = na + b\sum x_i \rightarrow ① \quad \& \quad \sum x_i y_i = a\sum x_i + b\sum x_i^2 \rightarrow ②$$

divide ① by n

$$\frac{\sum y_i}{n} = a + b\frac{\sum x_i}{n}$$

$$\Rightarrow a = \frac{\sum y_i}{n} - b\frac{\sum x_i}{n} \quad * \text{ sub into } ②$$

$$\sum x_i y_i = \frac{\sum x_i \sum y_i}{n} - b\frac{(\sum x_i)^2}{n} + b\sum x_i^2$$

$$\Downarrow$$

$$b = \frac{\sum x_i y_i - \sum x_i \sum y_i / n}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} * \frac{n}{n}$$

$$b = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

(2) Let $\{(x_i, y_i)\}_{i=1}^{n}$ be ranks individual distinct points

show that $r = \rho = 1 - \frac{6\sum d^2}{n(n^2-1)}$ where

$d_i = x_i - y_i$; $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$, $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

$$r = \frac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}}$$

"Proof"

Let $X = x - \mu_x$, $Y = y - \mu_y$

& x_i, y_i are distinct points then their ranks $1, 2, ---, n$

mean of their ranks are

$$\mu_{R_x} = \mu_{R_y} = \frac{1+2+ --- +n}{n} = \frac{\sum_{i=1}^{n} i}{n}$$

$$= \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{(n+1)}{2}$$

$$\because d = X - Y \Rightarrow \sum d^2 = \sum (X-Y)^2$$

$$= \sum X^2 - 2\sum XY + \sum Y^2$$

$$\sum X^2 = \sum (x - \mu_{R_x})^2 = \sum \left(x - \frac{n+1}{2}\right)^2$$

$$= \sum_{i=1}^{n} x^2 - (n+1) \sum_{i=1}^{n} x + \frac{(n+1)^2}{4} \, (n) \qquad \left(\sum_{i=1}^{n} 1 = n\right)$$

$$= \frac{n(n+1)(2n+1)}{6} - (n+1) \frac{n(n+1)}{2} + \frac{n(n+1)^2}{4}$$

$$= \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{4}$$

$$= \frac{n(n+1)}{12} \left[4n+2 - 3n-3 \right] = \frac{n(n+1)(n-1)}{12}$$

$$= \frac{n(n^2-1)}{12} = \sum Y^2$$

$$\therefore \sum d^2 = \frac{n(n^2-1)}{6} - \sum XY$$

$$\therefore \sum XY = \frac{1}{2} \left[\frac{n(n^2-1)}{6} - \sum d^2 \right]$$

$$\because \; r = \frac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}} = \frac{\sum XY}{\sum X^2} = \frac{\sum XY}{\sum Y^2}$$

$$= \frac{\frac{1}{2}\left[\frac{n(n^2-1)}{6} - \sum d^2\right]}{\frac{n(n^2-1)}{12}}$$

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} \quad \#$$

[3] Let μ_x and μ_y be the means of $\{x_i\}_{i=1}^{n}$ and $\{y_i\}$ respecively

and $X = x - \mu_x$, $Y = y - \mu_y$

show that $r = \frac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$

where $Cov(X,Y) = \frac{\sum XY}{n}$

"Proof"

$$\sum XY = \sum (x-\mu_x)(y-\mu_y)$$

$$= \sum xy - \mu_x \sum y - \mu_y \sum x + \mu_x \mu_y (n) \qquad \sum_{i=1}^{n} 1 = n$$

$$\mu_x = \frac{\sum x}{n} \;\&\; \mu_y = \frac{\sum y}{n}$$

$$= \sum xy - \frac{\sum x \sum y}{n} - \frac{\sum y \sum x}{n} + \frac{\sum y \sum x}{n}$$

$$= \sum xy - \frac{\sum x \sum y}{n} \rightarrow (1)$$

$\Sigma X^2 = \Sigma(x-\mu_x)^2$

where $\sigma_x^2 = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$ & $\sigma_y^2 = \frac{\Sigma y^2}{n} - \left(\frac{\Sigma y}{n}\right)^2 \rightarrow *$

& $\sigma^2 = \frac{\Sigma(x_i - \mu)^2}{n}$

$\therefore \Sigma X^2 = n\sigma_x^2 \rightarrow ②$ & $\Sigma Y^2 = n\sigma_y^2 \rightarrow ③$

sub ①, ②, ③

$\therefore r = \frac{\Sigma xy - \Sigma x \Sigma y / n}{n\sigma_x \sigma_y}$ → Cov(x, y) sub with *

$$r = \frac{\Sigma xy - \Sigma x \Sigma y / n}{n\sqrt{\frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n^2}}\sqrt{\frac{\Sigma y^2}{n} - \frac{(\Sigma y)^2}{n^2}}}$$

$$= \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}} \quad \#$$

4 show that $\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 - 2r\sigma_x\sigma_y$

"Proof"

$\because r = \frac{\Sigma XY}{n\sigma_x\sigma_y}$

$\because n\sigma_x^2 = \Sigma X^2 = \Sigma(x-\mu_x)^2$

$n\sigma_y^2 = \Sigma Y^2 = \Sigma(y-\mu_y)^2$

$\therefore n\sigma_{x-y}^2 = \Sigma(X-Y)^2 = \Sigma\left[(x-y) - \mu_{x-y}\right]^2$

where $\mu_{x-y} = \frac{\Sigma(x-y)}{n} = \frac{\Sigma x}{n} - \frac{\Sigma y}{n}$

$= \mu_x - \mu_y$

$\therefore n\sigma_{x-y}^2 = \Sigma\left[(x-\mu_x) - (y-\mu_y)\right]^2$

$$= \sum (x-\mu_x)^2 - 2\sum (x-\mu_x)(y-\mu_y) + \sum (y-\mu_y)^2$$

$$= n\sigma_x^2 + n\sigma_y^2 - 2nr\sigma_x\sigma_y$$

$$\therefore \sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 - 2r\sigma_x\sigma_y \quad \#$$

(5) show that the slope of regression line $y = a + \underline{b}x$ is $r\frac{\sigma_y}{\sigma_x} \rightarrow b = r\frac{\sigma_y}{\sigma_x}$??

"Proof"

$$y = a + bx \rightarrow *$$

$$\sum y_i = na + b\sum x_i \rightarrow (1)$$

$$\sum x_i y_i = a\sum x_i + b\sum x_i^2 \rightarrow (2)$$

$$\frac{(1)}{n} \Rightarrow \frac{\sum y_i}{n} = a + b\frac{\sum x_i}{n}$$

$$\mu_y = a + b\mu_x \rightarrow **$$

$$\& \quad r = \frac{\sum XY}{n\sigma_x\sigma_y} = \frac{Cov(x,y)}{\sigma_x\sigma_y}$$

$$\sum XY = \sum (x-\mu_x)(y-\mu_y) \rightarrow \text{sub } *, **$$

$$= \sum (x-\mu_x)(\cancel{a} + bx - \cancel{a} - b\mu_x)$$

$$= b\sum (x-\mu_x)^2 = nb\frac{\sum (x-\mu_x)^2}{n}$$

$$= nb\sigma_x^2$$

$$\& \quad nr\sigma_x\sigma_y = \sum XY$$

$$\cancel{n}r\cancel{\sigma_x}\sigma_y = \cancel{n}b\sigma_x^2 \rightarrow \boxed{b = r\frac{\sigma_y}{\sigma_x}} \quad \#$$